

**REMARKS**

In the Office Action, claims 1-43 were finally rejected under 35 U.S.C. § 102(e) as allegedly being anticipated by Application Publication No. 2001/0000962 of Rajan (hereinafter “Rajan”).

Applicants respectfully traverse the rejections of record, and further submit that claims 1-43 are in condition for allowance.

**Rejections under 35 U.S.C. § 102(e) in view of Rajan**

Claims 1-43 were rejected under 35 U.S.C. § 102(e) as allegedly anticipated by Rajan.

Independent claim 1 is directed to a system for generating a description record from multimedia information, comprising, *inter alia*:

a computer processor, coupled to said at least one multimedia information input interface, receiving said multimedia information therefrom, processing said multimedia information by performing *object extraction processing* to generate multimedia object descriptions from said multimedia information, and processing said generated multimedia object descriptions by *object hierarchy processing* to generate multimedia object hierarchy descriptions indicative of an organization of said object descriptions, wherein *at least one description record including said multimedia object descriptions and said multimedia object hierarchy descriptions is generated* for content embedded within said multimedia information

Independent claim 17 includes similar limitations in the context of a method claim.

Generally speaking, the present invention relates to MPEG-7, which includes techniques for describing and organizing multimedia information. As described in the

Background of the Invention (starting at p. 1 of the specification), the prior art provides means for searching textual information, both on the internet and locally. However, there was no means for searching multimedia content. An aim of MPEG-7 is to process multimedia such as video data to extract information about what is shown in the video and provide descriptions that may later aid in searching or cataloging the video.

Rajan is directed to a method and apparatus for composing and presenting multimedia programs using the MPEG-4 standard at a multimedia terminal, including an architecture wherein the composition of a multimedia scene and its presentation are processed by two different entities – a “composition engine” and a “presentation engine.” *See Rajan, ¶ 0002.* In this sense, the description of Rajan is limited to the MPEG-4 standard, “which allows a user to interact with video and audio objects within a scene,” and which allows a user to modify scenes by deleting, adding, or repositioning objects, or changing the characteristics of objects, such as size, color, and shape, for example. *See Rajan, ¶ 0004.* Rajan is thus directed to a different problem, i.e., composing and presenting multimedia video, from that of the present invention, which is instead directed to techniques for describing multimedia information content to enable intelligent searching of multimedia content via, e.g., the Internet. *See Specification, p.1 , lines 1-4, p. 9, lines 23-29.* This distinction is inherent in the differences between the field of the Rajan reference (MPEG-4) and the field of the present invention (MPEG-7), and would be immediately understood by one of ordinary skill in the art.

The Examiner, on pp. 4-6 of the Final Office Action, maintains that ¶¶ 0042 – 0046 of Rajan disclose all elements of claim 1. Applicants respectfully disagree. Not only do

these paragraphs *not* disclose all elements of the claim, they are not even directed to solving the same problem as the claimed invention, and, moreover, do not even relate to the same subject matter.

Rajan does not disclose or suggest “processing said multimedia information by *performing object extraction processing* to generate multimedia object descriptions.” The Examiner alleges that ¶ 0042 of Rajan discloses the claimed object extraction. However, ¶ 0042 of Rajan states:

According to the MPEG-4 Systems standard, the scene description information is coded into a binary format known as BIFS (Binary Format for Scene). This BIFS data is packetized and multiplexed at a transmission site, such as a cable and or satellite television headend, or a server in a computer network, before being sent over a communication channel to a terminal 100. The data may be sent to a single terminal or to a terminal population. Moreover, the data may be sent via an open-access network or via a subscriber network.

This portion of Rajan is directed to BIFS (Binary Format for Scene) coding, as set forth in the MPEG-4 standard. The BIFS data is packetized and transmitted over a communications link. Applicants do not see *any relation* between this paragraph and object extraction for generating multimedia descriptions, as claimed in the present invention.

Likewise, ¶ 0043 is cited as allegedly disclosing the claimed feature of “*processing said generated multimedia object descriptions by object hierarchy processing* to generate multimedia object hierarchy.” However, while ¶ 0043 generally discloses that an “MPEG-4 scene follows a hierarchical structure,” it nowhere indicates that multimedia

objection descriptions, which are generated by, e.g., performing object extraction, are processed by object hierarchy processing.

Furthermore, the claimed hierarchy and the hierarchy of Rajan are completely different. The cited paragraph of Rajan, ¶ 0043, states that “an MPEG-4 scene follows a hierarchical structure.” Notably, ¶ 0043 contains the *only* reference to hierarchy in Rajan. This hierarchy is not described with sufficient detail, and certainly is not enabling to one of ordinary skill in the art. For at least this reason, the claimed “processing said generated multimedia object descriptions by object hierarchy processing to generate multimedia object hierarchy” is not disclosed for purposes of anticipation.

Moreover, the claimed hierarchy entirely different from that of the present invention. The cited paragraph of Rajan, ¶ 0043, states that “an MPEG-4 scene follows a hierarchical structure.” This would suggest that the “tree structure” discussed in Rajan is for purposes of flow of a video scene in space and time. This is further apparent when considered in the context of the problem which Rajan is directed to, i.e., *composing* and *presenting* multimedia video *in space and time*. However, the claimed hierarchy, as further described, e.g., at p. 13 of the present application, relates to an object hierarchy *for description of particular video objects with varying levels of specificity* – for purposes of *content description*, and *not* hierarchy of a scene for composing or presenting the scene. In addition, the hierarchy of the present invention is not only based on temporal and spatial constraints, but may also include visual, motion and semantic information. This is further apparent by reference to a comparison of MPEG-4 and MPEG-7 standards papers. For the Examiner’s reference, Applicants previously directed the Examiner to the following two

hyperlinks to MPEG standards papers:

<http://www.chiariglione.org/mpeg/standards/mpeg-7/mpeg-7.htm>

<http://www.chiariglione.org/mpeg/standards/mpeg-4/mpeg-4.htm>

which provide further detail regarding MPEG-4, MPEG-7, and the differences in the hierarchical structures of each.

However, with respect to the citations to the MPEG standards available at the links above, the Examiner stated in the Final Office Action that:

“The above constitutes new matter and will not be entered into the record. Furthermore, above disclosure is objected to because it contains an embedded hyperlink and/or other form of browser-executable code. Applicant is required to delete the embedded hyperlink and/or other form of browser-executable code. See MPEP § 608.01”

This objection is improper for a myriad of reasons. First, the hyperlinked information was not provided in the specification of the application, and accordingly cannot possibly be “new matter.” The MPEP section cited by the Examiner applies only where an applicant includes such information in a patent application specification. Applicants here have not. Rather, Applicants have provided this information to aid the Examiner in reviewing the prior art and to demonstrate the substantial differences between the claimed invention and the prior art. Accordingly, there is no basis for nor reason to “delete the embedded hyperlink,” as the Examiner requires in the Final Office Action on page 9. For the Examiner’s convenience, paper copies of the documents provided at the above-cited hyperlinks are enclosed herewith as Exhibits A and B to this Response. Applicants again urge the Examiner to review these papers for additional explanation of the differences

between the present invention, which, as would be understood by one of ordinary skill in the art, relates to the field of MPEG-7 technology, and the cited reference, which relates to MPEG-4 technology.

Accordingly, because Rajan fails to disclose or suggest at least these claimed features, and is in fact related to an entirely different object and invention, this reference fails to anticipate independent claims 1 and 17. Applicants respectfully submit that these claims are in condition for allowance. Additionally, because all depending claims 2-16 and 18-32 contain the foregoing limitations through dependency from claims 1 and 17, Applicants respectfully submit that these claims are also in condition for allowance.

Claims 3, 7, 19 and 23 are patentable for the additional reason that they include limitations of “image segmentation” and “feature extraction.” The Examiner asserts that ¶ 0045 of Rajan, copied below, describes the claimed features:

The scene description information can also indicate attribute value selection. Individual media objects and scene description nodes expose a set of parameters to a composition layer through which part of their behavior can be controlled. Examples include the pitch of a sound, the color for a synthetic object, activation or deactivation of enhancement information for scaleable coding, and so forth.

Rajan, ¶ 0045.

Applicants again cannot find any reference in the above-cited paragraph to the claimed “image segmentation” and “feature extraction.” Indeed, there is none. Accordingly, for at least this additional reason, claims 3, 7, 19 and 23 are further patentable over Rajan.

Independent claim 33 is directed to a computer readable medium with at least one multimedia description record describing multimedia content for corresponding

multimedia information, the description record comprising, *inter alia*, object descriptions, generated by performing object extraction processing. The Examiner has rejected claim 33 for the same reasons as claims 1 and 17. However, as discussed at length above, Rajan fails to disclose or even remotely suggest at least several features of claim 33, including the feature of object descriptions generated by performing object extraction processing. Accordingly, because Rajan fails to disclose or suggest at least these claimed features, this reference fails to anticipate independent claim 33. Applicants respectfully submit that this claim is in condition for allowance. Additionally, because all depending claims 34-43 contain the foregoing limitations through dependency, Applicants respectfully submit that these claims are also in condition for allowance.

Additionally, there are other deficiencies in the Examiner's "Response to Arguments," which begins at p. 4 of the Final Office Action.

First, on page 8., the Examiner states the following:

"Examiner is not persuaded. In response to applicant's argument that the references fail to show certain features of applicant's invention, it is noted that the features upon which applicant relies (i.e., generation of multimedia object hierarchy) is not recited in the rejected claim(s)."

However, Applicants refer to text in, e.g., claim 1, which states: "processing said generated multimedia object descriptions by object hierarchy processing to generate multimedia object hierarchy descriptions."

In this respect, Rajan and MPEG-4 generally deals with hierarchy in the sense of *scene* descriptions, wherein objects are described in a hierarchical fashion to provide composition and presentation information for presenting multimedia content. The present

invention, however, refers to object hierarchy descriptions, as described throughout the specification, including at, e.g., pp. 19-20. The descriptions include information which is useful for searching a library of multimedia segments, such as “names of the picture, the names of persons in the picture, the location where the picture was taken, the event that is represented by the picture, the date of the picture, color features... .” (Specification, p. 20). The distinctions between the types of hierarchical structures claimed and those mentioned briefly in Rajan would be apparent to one of ordinary skill in the art, as the MPEG-4 and MPEG-7 standards are completely different, have different structures, and operate to achieve different goals. For at least these additional reasons, Applicants respectfully submit that the rejections in the Final Office Action should be withdrawn.

Additionally, the Examiner relies on a 1999 MPEG-4 publication as alleged evidence of the state of the art at the time the present application was filed. (*See* Final Office Action, p. 9) However, the present application claims priority to November 6, 1998. For at least this reason, the cited reference is not a valid indicator of the state of the art at the time the present application was filed. Additionally, as explained above, the cited reference refers to MPEG-4 technology, which is distinct from the claimed invention. For at least these additional reasons, Applicants respectfully submit that the rejections in the Final Office Action should be withdrawn.

Finally, on p. 10 of the Final Office Action, the Examiner refers to a portion of Rajan which allegedly “reads on image segmentation and feature extraction”:

4. The MPEG-4 communication standard allows a user to interact with video and audio objects within a scene, whether they are from conventional sources, such as moving video, or from synthetic (computer generated) sources. The user can

modify scenes by deleting, adding or repositioning objects or changing the characteristics of the objects, such as size, color, and shape, for example.

6. The objects can exist independently, or be joined with other objects in a scene in a grouping known as a “composition.” Visual objects in a scene are given a position in two- or three- dimensional space while audio objects can be placed in a sound space.

8. BIFS commands can add or delete objects from a scene, for example, or changed [sic] the visual or acoustic properties of objects. BIFS commands also define, update, and position the objects. For example, a visual property such as the color or size of an object can be changed, or the object can be animated.

None of these cited portions of Rajan discloses or even remotely suggests *image segmentation* or *feature extraction* – which is entirely expected, since the above reference deals with MPEG-4 (intended for composing and presenting multimedia content) and is *not* directed to MPEG-7 (intended for extracting information from multimedia content in order to categorize and search the multimedia content). As discussed above, for at least this additional reason, Applicants respectfully submits that the rejections in the Final Office Action should be withdrawn.

**CONCLUSION**

In view of the foregoing remarks, favorable consideration and allowance of claims 1-43 are respectfully solicited. In the event that the application is not deemed in condition for allowance, the examiner is invited to contact the undersigned in an effort to advance the prosecution of this application.

Respectfully submitted,



Paul A. Ragusa  
PTO Reg. No. 38,587

Robert L. Maier  
PTO Reg. No. 54,291

*Attorneys for Applicants*  
(212) 408-2538

BAKER BOTTS L.L.P.  
30 Rockefeller Plaza  
New York, NY 10112